

**Nevoeiro ou Neblina no Amanhecer: Variação Linguística nos Dados do Atlas
Linguístico do Brasil (ALiB)**

Fog or Mist at Dawn: Linguistic Variation in the Linguistic Atlas of Brazil (ALiB)

Genivaldo da Conceição Oliveira

Universidade Federal do Recôncavo da Bahia – UFRB

Resumo: Este trabalho apresenta os resultados da análise do léxico registrado, nos estados da Bahia e do Paraná, no campo semântico *fenômenos atmosféricos* com o objetivo de contribuir para um melhor conhecimento do Português Brasileiro, tal como se apresenta nas cidades que constituem a rede de pontos do Atlas Linguístico do Brasil (ALiB) nesses dois estados. Nossa pesquisa se baseia nos princípios teóricos da Dialetoлогия, Sociolinguística e Lexicologia. O *corpus* está conformado por um extrato dos dados do ALiB constituído das perguntas 21 e 22 do Questionário Semântico-Lexical (QSL) e se utiliza do método da Geolinguística para a análise espacial dos dados. Enfatizamos o aspecto diatópico, contudo, recorremos, de maneira periférica, à análise de outras variáveis como a diastrática e a diageracional. Concluímos que para as questões 21 e 22, as variantes mais produtivas *neblina* e *amanhecer* aparecem distribuídas por todas as mesorregiões da Bahia e do Paraná. As variantes que ocorreram apenas na Bahia ou apenas no Paraná não foram fornecidas por um número expressivo de informantes. Verificamos, portanto, que no plano da análise diatópica, nesse recorte analisado, como resultado mais relevante o fato de a distribuição das variantes terem um caráter mais homogêneo do que heterogêneo.

Palavras-chave: Dialetoлогия. Geolinguística. Sociolinguística. Variação.

Abstract: This paper presents the results of the analysis of the registered lexicon, in the states of Bahia and Paraná, in the semantic field *atmospheric phenomena* with the objective of contributing to better knowledge of the Brazilian Portuguese, as it is presented in the cities that constitute the network of points of the Linguistic Atlas of Brazil (ALiB) in these two states. Our research is based on the theoretical principles of Dialectology, Sociolinguistics and Lexicology. The *corpus* is made up of an extract of ALiB data consisting of questions 21 and 22 of the Semantic-Lexical Questionnaire (QSL) and uses the Geolinguistic method for the spatial analysis of the data. We emphasize the diatopic aspect, however, we resort, in a peripheral way, to the analysis of other variables such as diastratic and deagerational. We conclude that for questions 21 and 22, the most productive *fog* and *dawn* variants appear distributed throughout the mesoregions of Bahia and Paraná. The variants that occurred only in Bahia or only in Paraná were not provided by a significant number of informants. We verified, therefore, that in terms of diatopic analysis, in this analyzed section, the most relevant result is the fact that the distribution of variants has a more homogeneous character than heterogenuous.

Key-words: Dialectology; Geolinguistics; Sociolinguistics; Variation.

Recebido em 25 de junho de 2023

Aprovado em 20 de dezembro de 2023.

Introdução

Descoberto em 1500, o Brasil só começou propriamente a ser colonizado em 1534 com o regime das capitanias hereditárias. A partir desta data tem-se a efetiva introdução da língua portuguesa.

Como é sabido e assinalam diferentes autores, a língua portuguesa no Brasil, no curso da sua história, vem passando por muitas modificações, realidade para a qual, nos meados do século XX, já chamava a atenção de Nascentes (1953, p. 9-10) ao registrar que:

A língua portuguesa no Brasil sofreu grandes alterações porque teve de ser aprendida por homens de duas raças que falavam línguas de estrutura inteiramente diversa do tipo flexional. O índio foi o primeiro que aprendeu o português; é natural, pois foi o povo autóctone. Só mais tarde aparece o outro fator etnográfico, o negro. [...] A escravidão vermelha precedeu à negra e daquela já se fala em 1531, quando Martim Afonso concedeu a Pedro de Góis permissão de levar para Europa dezessete escravos índios; mas desde cedo na capitania de S. Vicente são escravos negros que trabalham na agricultura da cana.

Este estudo investiga a relação entre o léxico referente à área semântica *fenômenos atmosféricos* documentada no estado da Bahia e no estado do Paraná, com base no *corpus* do Atlas Linguístico do Brasil (ALiB) a partir de dados de 164 informantes, distribuídos: 16 nas duas capitais, 84 nas cidades do interior da Bahia e 64 nas cidades do interior do Paraná. Nesta pesquisa, analisamos as respostas das questões 21 e 22 do Questionário Semântico-Lexical (QSL) do ALiB. Identificamos e estabelecemos as semelhanças e diferenças encontradas neste recorte estudado nos dois estados com base teórica na abordagem apresentada pela Geolinguística. Utilizamos, também, pressupostos teóricos da Dialetoлогия, Sociolinguística e Lexicologia. Analisamos esse recorte do ALiB entre a Bahia e o Paraná, dois estados brasileiros que não apresentam contiguidade geográfica e exibem tipos de povoamento diferentes, para observar se o léxico fornecido pelos informantes apresenta mais homogeneidade ou heterogeneidade.

1. Dialetoлогия e seu percurso histórico – uma breve análise

A Dialetoлогия identifica, situa e descreve os usos diferentes em que a língua varia de acordo com sua disposição espacial, histórica e sociocultural, respondendo a um pensamento mais amplo, pois, como afirma Cardoso (2010, p. 27),

O interesse pelo estudo da diversidade de usos da língua e a evidência de certa preocupação universal com as diferenças dialetais perpassam a história dos povos em todos os momentos, ora como simples constatação, ora como instrumento político, ora como mecanismo de descrição das línguas.

A Dialetoлогия é, como a própria palavra sugere, o estudo dos dialetos. Na linguagem coloquial, diz-se que um dialeto é uma língua não padrão, que geralmente se associa a grupos que não ostentam prestígio ou uma língua que não tem tradição escrita. Entretanto, Chambers e Trudgill (1994, p. 19) não aprovam estes conceitos e dizem que “todos os falantes são falantes de, pelo menos, um dialeto”¹ e que a forma padrão de uma língua constitui-se em si um dialeto. Esta definição, contudo, esbarra em um problema: Como distinguir língua de dialeto? Estes autores afirmam que uma língua é um conjunto de dialetos mutuamente inteligíveis, embora salientem que tal definição não seja totalmente satisfatória porque esta tal inteligibilidade nos traz alguns problemas. Eles citam as línguas norueguesa, sueca e dinamarquesa como três línguas distintas, mas mutuamente inteligíveis. Em contrapartida, a língua alemã, considerada como uma única língua, apresenta problemas de comunicação entre os diferentes falantes do alemão, usuários de distintos dialetos. A inteligibilidade entre as línguas apresenta graus para mais ou para menos e pode não se apresentar igual nas duas direções.

Chambers e Trudgill (1994) também traçam um paralelo entre o conceito de dialeto e sotaque. Para eles, sotaque tem a ver com a forma que um falante pronuncia a língua e, portanto se aplica à variedade fonética ou fonológica da língua. Por outro lado, dialeto, refere-se às variedades que são diferentes do ponto de vista gramatical ou lexical, além do fonológico. Contudo, eles esclarecem que os sotaques e os dialetos frequentemente se fundem uns com os outros sem que tenham uma separação definida. Estes autores afirmam que, embora estas observações acerca de diferenças dialetais sejam muito comuns, o estudo dos dialetos só começa de maneira sistemática na

¹ “... todos los hablantes lo son al menos de un dialecto...”. Tradução nossa.

segunda metade do século XIX. Finch (2000, p. 215)² afirma que um “dialeto é uma variedade linguística restrita a um espaço geográfico com formas sintáticas e itens vocabulares distintos”. Para ele, geralmente, “distingue-se de sotaque; o qual se refere apenas a aspectos de pronúncia, embora em algumas ocasiões, o dialeto vagamente inclui o sotaque”. Muitos dialetos são regionais em sua origem e pertencem a uma área específica. Como podemos ver, há confluência de ideias sobre a acepção de dialeto entre Finch e Chambers e Trudgill. Finch acrescenta que a Dialetologia mudou seu foco do estudo sobre dialetos tradicionais para o estudo dos dialetos modernos, urbanos. Segundo ele, um dos nomes mais importantes tem sido o do sociolinguista americano William Labov, cujo trabalho inicial sobre a fala de novaiorquinos influenciou uma geração de sociolinguistas.

A Dialetologia era vista nos seus primórdios sob uma perspectiva preponderantemente diatópica. Os primeiros estudos dialetológicos eram predominantemente focalizados dentro de um espaço e tinham uma abordagem monodimensional. Aos poucos, a monodimensionalidade foi perdendo sua hegemonia para um estudo mais pluridimensional. De acordo com Cardoso (2010, p. 15), a “Dialetologia é um ramo dos estudos linguísticos que tem por tarefa identificar, descrever e situar os diferentes usos em que uma língua se diversifica, conforme a sua distribuição espacial, sociocultural e cronológica”. O estudo dialetológico obedece a três passos importantes: identificar, descrever e situar a variação linguística. Após a realização do primeiro passo, que é a identificação do fenômeno linguístico, passa-se a descrevê-lo fazendo o levantamento das variantes. Descrever é enumerar as variantes lexicais possíveis e que tenham o mesmo valor de verdade, como em *macaxeira*, *aipim* e *mandioca*. A Dialetologia inventaria, sistematiza e descreve estas variações.

Chambers e Trudgill (1994, p. 139) afirmam que uma das maiores preocupações da Dialetologia tradicional ou Geografia Linguística tem sido a determinação de isoglossas, dos limites entre duas regiões que diferem em algum traço linguístico entre si. Ao analisar o significado literal de isoglossa “iso → igual” e “glossa → língua”, eles observam que isoglossa, supostamente, “quer expressar o fato de que uma linha traçada

² “Dialect is a geographically based language variety with distinct syntactic forms and vocabulary items”. “It’s usually distinguished from accent, which refers solely to features of pronunciation, although on occasions dialect is loosely used to include accent”. Tradução nossa.

através de uma região mostrará duas áreas em cada uma das quais coincide algum aspecto do uso linguístico, mas que difere uma da outra”.³

Ferreira e Cardoso (1994) observam que um feixe de isoglossas demarca um dialeto. É, portanto, um conjunto de isoglossas que se somam e exibem uma relativa homogeneidade dentro de uma comunidade linguística em confronto com outras. Elas adicionam que, devido a esta relativa homogeneidade, podemos crer que não há limites rígidos entre as línguas, uma vez que toda língua histórica é constituída por um conjunto de dialetos. Isto corrobora o pensamento de Chambers e Trudgill.

Com base em Cardoso (2010), observamos que há duas características importantes na origem da Dialectologia independentemente do princípio metodológico usado. A primeira característica é o reconhecimento das diferenças ou das semelhanças que a língua transmite. Outra característica é o estabelecimento das relações entre as diversas manifestações linguísticas documentadas ou entre elas e a ausência de dados registrados, circunscritos a espaços e realidades prefixados. A Dialectologia não pode desconsiderar fatores extralinguísticos, próprios do falante, da mesma maneira que não pode desconsiderar as implicações que estes fatores acarretam nos atos da fala. De maneira que idade, sexo, escolaridade e características socioculturais se tornam elementos de pesquisa que convivem com a busca de identificação de áreas dialetais. Neste ponto, é possível ver uma confluência de propósitos entre a Dialectologia e a Sociolinguística uma vez que ambas as disciplinas estudam a variação linguística. Portanto, os enfoques diatópico e sociolinguístico estão presentes tanto na Dialectologia quanto na Sociolinguística. Todavia, o que as distingue é a forma de tratar os fenômenos e a perspectiva que cada uma imprime à abordagem dos fatos linguísticos. A Dialectologia tem como base da sua descrição a localização espacial dos fatos estudados, demonstrando seu caráter eminentemente diatópico, embora considere fatores sociais.

No Brasil, a história da Dialectologia é marcada pela presença de estudiosos como Amadeu Amaral e Antenor Nascentes que publicaram alguns dos primeiros trabalhos sobre a Dialectologia brasileira. Ferreira e Cardoso (1994, p. 37), tomando por base a divisão em duas fases realizada por Nascentes, observam que “podemos dividir a história dos estudos dialetais em três grandes fases”.

³ “Presumiblemente quiere expresar el hecho de que una línea trazada a través de una región mostrará dos áreas em cada una de las cuales coincide algún aspecto del uso linguístico, pero que difiere una de la otra”. Tradução nossa.

A primeira fase compreende o período de 1826 a 1920, que culmina com a publicação de *O dialeto caipira* de Amadeu Amaral. A tônica deste período foi o estudo do léxico relativo ao português do Brasil e a criação de vários dicionários, vocabulários e léxicos de caráter regional. Ferreira e Cardoso (1994) afirmam que a segunda fase se inicia com a publicação de *O dialeto Caipira* de Amadeu Amaral. Neste período, há um grande número de trabalhos que tratam da gramática, embora haja também vários de cunho lexicográfico. Desta segunda fase, podemos destacar dois trabalhos iniciais: *O dialeto caipira* em 1920, como já mencionamos, e *O linguajar carioca* de Antenor Nascentes em 1922.

Em *O dialeto caipira*, Amadeu Amaral chama a atenção para a pesquisa *in loco* para dissipar falsas hipóteses e conclusões que não refletiam a verdadeira realidade linguística. Ele traça, para futuros pesquisadores da Dialectologia, passos básicos para um trabalho mais criterioso: a observação imparcial, sistemática no trabalho, a retratação fiel da realidade a partir do que as amostras coletadas permitiam e a verificação pessoal dos fatos para eliminar tudo que fosse hipotético e incerto. Amadeu Amaral abriu, assim, o caminho para os estudos dialetais no Brasil com linhas gerais para um estudo monográfico de uma região. Antenor Nascentes, em *O linguajar carioca*, começa traçando linhas gerais para a compreensão do português brasileiro que ele chama de *o falar brasileiro* e situa o linguajar carioca neste grupo. Para Nascentes (1953), a enorme extensão territorial do Brasil, sem fáceis meios de comunicação interior, quebrou a unidade do falar, fragmentando-o em subfalares. Esta fragmentação é também influenciada pelo modo diferente de povoamento de cada região. É palpável a diferença entre a fala *cantada* do nortista e a fala *descansada* do sulista. A estas obras *O dialeto caipira* e *O linguajar carioca*, podemos adicionar outras nesta segunda fase, como *O vocabulário pernambucano* de Pereira da Costa (1976), que enfocam o léxico regional seguindo a linha dominante na fase anterior.

A terceira fase se inicia em 1952 e se distingue pela elaboração de trabalhos baseados em *corpus* constituído de forma sistemática e surge, então, neste momento, a preocupação com a execução e desenvolvimento dos estudos da Geolinguística no Brasil e com a produção de um atlas linguístico do Brasil. Nesse sentido, o governo brasileiro toma a iniciativa de atribuir à Comissão de Filologia, da recém-criada Casa Ruy Barbosa, através do decreto 30.643 a responsabilidade pela produção do atlas linguístico do Brasil. A terceira fase da história dos estudos dialetais tem, assim, como

marca identificadora, o começo dos estudos sistemáticos no campo da Geografia Linguística. Entretanto, Mota (2006) complementa a proposta de Ferreira e Cardoso (1994) com uma quarta fase. Esta quarta fase começa com a retomada do Projeto Atlas Linguístico do Brasil em 1996. Essa nova fase coincide com a incorporação dos princípios implementados pela Sociolinguística, a partir da década de 60 do século XX, abandonando-se a visão monodimensional que era predominante na Geolinguística que atualmente chamamos de tradicional. Cardoso (2009) observa que é importante reiterar que a implantação do Projeto Atlas Linguístico do Brasil em 1996 é o marco mais significativo para estabelecer esta quarta fase dos estudos dialetológicos, atrelados ao estudo da variação linguística, que transcende limites geográficos e está presente em todas as comunidades de fala.

2. Variação linguística

Formas linguísticas em variação estão presentes em todas as comunidades de fala. Estas formas são chamadas de variantes que são, na verdade, maneiras diferentes de falar a mesma coisa no mesmo contexto e com o mesmo valor de verdade. Estas variantes, por sua vez, estão sempre competindo dentro da comunidade de fala à qual pertencem. Desta maneira, temos as variantes padrão e não padrão, aquelas que são conservadoras contra as que são inovadoras e as variantes que recebem algum tipo de estigma em oposição àquelas de prestígio. Geralmente, uma variante padrão é considerada conservadora e possui maior importância sociolinguística dentro da comunidade. Em contrapartida, uma variante inovadora tende ser não-padrão e é, portanto, estigmatizada pelos falantes da comunidade a que pertence. A título de ilustração, trazemos a presença do segmento fônico /s/ como marca de plural no sintagma nominal que é a forma padrão, conservadora e, portanto, de prestígio. Ao passo que a não marcação do plural /s/ no sintagma nominal é estigmatizada.

Labov (2008, p. 260) observa que “no curso da evolução linguística, a mudança caminha para se completar, e regras variáveis se tornam invariantes. Quando isso acontece, há outra mudança estrutural que compensa a perda de informação envolvida”. Isto significa dizer que se uma regra variável for constante, ela oferece aos aprendizes da língua informação suficiente para manter as distinções básicas e as formas subjacentes. Portanto, nem tudo que varia sofre mudança, mas toda mudança linguística pressupõe variação uma vez que mudança é variação. Para Coseriu (1979, p. 64), a

língua não muda completamente porque se refaz. O falante não cria integralmente a sua expressão, mas utiliza o sistema que lhe é oferecido pela comunidade, além disso, aceita também a realização que a norma tradicional lhe fornece. Ele não inventa totalmente sua expressão, mas utiliza modelos anteriores porque este indivíduo é um ser histórico e porque a língua pertence a sua historicidade. Isto quer dizer que a expressão que é usada pelo falante tem uma história que a precede.

A diversidade linguística está relacionada não apenas com a questão territorial, mas também com a questão da desigualdade social. Bagno (2000) observa que:

No Brasil, embora a língua falada pela grande maioria da população seja o português, esse português apresenta um alto grau de diversidade e de variabilidade, não só por causa da grande extensão territorial do país – que gera as diferenças regionais, bastante conhecidas e também vítimas, algumas delas, de muito preconceito – mas principalmente por causa da trágica injustiça social que faz do Brasil o segundo país com a pior distribuição de renda em todo mundo. São essas graves diferenças de status social que explicam a existência, em nosso país, de um verdadeiro abismo linguístico entre falantes das variedades não-padrão de português brasileiro – que são a maioria de nossa população – e os falantes da (suposta) variedade culta, em geral mal definida, que é a língua ensinada na escola. (BAGNO, 2000, p. 16).

Idade, sexo, raça (ou cultura), profissão, posição social, grau de escolaridade são alguns dos fatores que desencadeiam variedade linguística. Para Preti (2003), as variantes decorrentes de faixas etárias, considerando o locutor adulto, restringem-se mais ao vocabulário. Para ele, a chamada linguagem jovem se refere a um vocabulário gírio, cujos limites são meio vagos. Este autor diz que a oposição entre a linguagem do homem e a linguagem da mulher pode indicar diferenças evidentes, sobretudo no léxico por conta de tabus morais. Contudo, ele aponta que essa oposição vem perdendo sua significação, especialmente nas cidades grandes, porque os meios de comunicação de massa, o teatro, a transformação dos costumes e padrões morais têm exercido um papel nivelador expressivo. Outro fator é a profissão do indivíduo que funciona no campo da linguagem técnica em que os falantes usam um vocabulário condizente com sua atividade. Além disso, a posição social requer que o falante tenha um cuidado especial com a linguagem que usa visando ter destaque dentro do grupo em que atua.

No momento em que nos referimos aos fatores que ocasionam o surgimento de variantes linguísticas, apontamos sempre que tais fenômenos ocorrem dentro de uma comunidade de fala, que é assim conceituada por Moreno Fernandez (1998):

Uma comunidade de fala é formada por um conjunto de falantes que compartilham efetivamente, pelo menos, uma língua, contudo, além disso, compartilham um conjunto de normas e valores de natureza sociolinguística: compartilham as mesmas atitudes linguísticas, as mesmas regras de uso, um mesmo critério na hora de valorizar socialmente os fatos linguísticos, os mesmos padrões sociolinguísticos. Os membros de uma comunidade de fala são capazes de se reconhecerem quando compartilham opiniões sobre o que é vulgar, o que é familiar, o que é incorreto, o que é arcaizante ou antiquado. Por isso, o cumprimento das normas sociolinguísticas que obriga o pertencimento a uma comunidade pode servir de marca diferenciadora, de marca de grupo, e por isto os membros de uma comunidade costumam acomodar seu discurso a normas e valores compartilhados. Uma comunidade de fala é basicamente uma comunidade de consenso, de sintonia entre grupos e indivíduos diferentes, onde conflitos são minimizados. (MORENO FERNÁNDEZ, 1998, p. 19-20).⁴

A variação linguística acontece em diversos níveis da fala dentre os quais a variação fonética e fonológica parece ser a mais estudada e, portanto, mais conhecida. As variantes de um fonema, geralmente, não supõem nenhuma mudança de significado. A variação gramatical, ou seja, morfológica e sintática, tal qual a fonético-fonológica, pode ocorrer por conta de fatores linguísticos ou pela combinação de fatores linguísticos e sociais. Como é autoexplicativo, as variáveis morfológicas afetam elementos da morfologia, cuja variação raras vezes implica níveis sintáticos e pragmáticos e que costumam ser determinadas por fatores tanto sociolinguísticos e estilísticos quanto por fatores históricos e geográficos.

Uma das dificuldades para o estudo da variação lexical é a determinação de correspondência entre variantes. Esta análise encontra problemas, especialmente no nível semântico-lexical, quanto à existência ou impossibilidade de explicar as equivalências por um viés teórico da sinonímia. Moreno Fernández (1998, p 29) observa que “a Sociolinguística tem se tornado, quase por necessidade epistemológica, uma defensora da existência da sinonímia, pelo menos no nível do discurso”. As unidades léxicas podem se encontrar semanticamente neutralizadas no discurso. Contudo, há dificuldade em demonstrar que duas ou mais variantes são equivalentes. Por conta da escassa frequência com que variantes lexicais alternam no discurso, alguns

⁴ “Una comunidad de habla está formada por un conjunto de hablantes que comparten efectivamente, al menos, una lengua, pero que, además, comparten unas mismas actitudes lingüísticas, unas mismas reglas de uso, un mismo criterio a la hora de valorar socialmente los hechos lingüísticos, unos mismos patrones sociolingüísticos[...] Los miembros de una comunicadad de habla son capaces de reconorcerse cuando comparten opinión sobre lo que es vulgar, lo que es familiar, lo que es incorrecto, lo que es arcaizante o anticuado. Por eso el cumplimiento de las normas sociolingüísticas al que obliga la pertenencia a una comunidad puede servir de marca diferenciadora, de marca de grupo, y por eso los miembros de una comunidad suelen acomodar su discurso a las normas y valores compartidos[...] Una comunidad de habla es básicamente una comunidad de consenso, de sintonia entre grupos e individuos diferentes, donde el conflicto está minimizado.” Tradução nossa.

estudiosos têm optado por localizar variáveis fora do discurso natural e continuado por meio de pesquisas e questionários.

3. Geografia Linguística

A Geografia Linguística é uma metodologia, ou seja, um conjunto de métodos para compilar de maneira sistemática as demonstrações das diferenças dialetais. A Geografia Linguística (Geolinguística) procura criar uma base empírica sobre a qual se possam extrair conclusões a respeito da variedade linguística que ocorre em um lugar determinado. A Geolinguística revelou uma heterogeneidade que não se concebia antes e, portanto, jogava por terra toda e qualquer pressuposição de ausência de exceções.

Segundo Cardoso (1998), a Geolinguística no Brasil ganha corpo em meados do século XX quando surgem as primeiras manifestações pela produção de um atlas linguístico do Brasil. A partir deste ponto, a pesquisa no campo da Dialetoлогия tem se desenvolvido. Estas pesquisas não caminharam em direção à elaboração de um atlas de abrangência nacional, mas buscavam mostrar realidades regionais, com diversos trabalhos publicados em vários pontos do país, efetivando atlas linguísticos por região. O primeiro Atlas publicado em território brasileiro foi o *Atlas Prévio dos Falares Baianos – APFB*, em 1963 de autoria de Nelson Rossi, Carlota Ferreira e Dinah Isensee.

4. Atlas Linguístico do Brasil (ALiB)

Em 1996, retoma-se a ideia de construção de um atlas nacional com o surgimento do Projeto Atlas Linguístico do Brasil – ALiB, durante o *Seminário Nacional Caminhos e Perspectivas para a Geolinguística no Brasil*, realizado em Salvador, na Universidade Federal da Bahia (UFBA), quando se constituiu um comitê nacional para elaboração do atlas. Este comitê, presidido pela Professora Doutora Suzana Cardoso da UFBA, conta com a participação de autores de atlas já publicados e em andamento. O ALiB documenta dados linguísticos no Brasil de Norte ao Sul, do Leste ao Oeste e descreve a realidade linguística do Português Brasileiro enfatizando a identificação das diferenças diatópicas, que podem ser fônicas, morfossintáticas e léxico-semânticas dentro da perspectiva da Geolinguística. São 250 pontos que constituem sua rede de localidades que reúnem 1.100 informantes documentados. O informante tem perfil que está atrelado ao espaço em que ele vive.

5. Bahia e Paraná

A Bahia se localiza no sul da região nordeste do Brasil, limitando-se ao Leste pelo oceano Atlântico, ao Norte com Sergipe, Alagoas, Pernambuco e Piauí, ao Sul com os estados de Minas Gerais e Espírito Santo e Oeste com Goiás e Tocantins.⁵

Segundo Reis (2009, p. 19), a província da Bahia era uma das mais prósperas regiões canavieiras das Américas no século XIX. Os engenhos de açúcar, puxados por mão-de-obra escrava, estavam situados especialmente no Recôncavo, região fértil e úmida que envolve a Baía de Todos os Santos. Reis afirma que Salvador, que naquele então era mais conhecida como Cidade da Bahia, ocupa um dos extremos desse conjunto geográfico. O território baiano começou a se estruturar pela faixa costeira, ainda no século XVI, partindo da cidade de Salvador e das vilas de Porto Seguro e de São Jorge dos Ilhéus. Em seguida, começaram a surgir outros núcleos populacionais em torno dos engenhos de açúcar e de pequenas propriedades de criação de gado.

O Paraná se localiza no Sul do Brasil, limitando-se ao Norte com o estado de São Paulo, a Leste com o oceano Atlântico, ao Sul com o estado de Santa Catarina e a Oeste com o estado do Mato Grosso e com as repúblicas do Paraguai e Argentina.⁶ Aguilera (2002) afirma que a história do Paraná compreende a composição de três comunidades regionais:

O Paraná Tradicional, que se esboçou no século XVII, com a procura do ouro, e se estruturou no século XVII sobre o latifúndio campeiro dos Campos Gerais, com base na criação e comércio do gado e, mais tarde, nas atividades extrativistas e no comércio exportador da erva-mate e da madeira, e as do Paraná Moderno, já no século XX, sendo as do Norte, com a agricultura tropical do café e que, a princípio, pelas origens e interesses históricos, ficou mais diretamente ligada a São Paulo, e a do Sudoeste e Oeste, dos criadores de suínos e plantadores de cereais que, pelas origens e interesses históricos, ficou a princípio mais intimamente ligada ao Rio Grande do Sul. (AGUILERA, 2002, p. 19).

Durante os séculos XVI e XVII, houve uma disputa pela posse do território paranaense pelas missões jesuíticas espanholas e pelas bandeiras paulistas sob ordens do governo português. Durante estes dois séculos, vários núcleos de povoamento começaram a surgir ao longo dos principais rios.

⁵ Disponível em [<https://www.sogeografia.com.br/Conteudos/Estados/Bahia/>] Acesso em 25 de maio de 2023.

⁶ Atlas Linguístico do Paraná. Vol. I. p. 21

6. Metodologia

Este trabalho se fundamenta na metodologia e no *corpus* do Projeto Atlas Linguístico do Brasil (ALiB) no que tange à área semântica *fenômenos atmosféricos* do Questionário Semântico-lexical (QSL – COMITÊ Nacional, 2001). Neste estudo priorizamos a variação diatópica, mas, seguindo os passos da Geolinguística Pluridimensional, também consideramos aspectos relativos às variações diastrática, diageracional e diassexual.

6.1. Corpus

O *corpus* está constituído pelas respostas às questões 21 e 22 do Questionário Semântico Lexical do Atlas Linguístico do Brasil aplicado em 22 cidades na Bahia e 17 no Paraná.

6.2. Localidades da Bahia

As localidades da Bahia, que perfazem o total de 22 pontos, estão distribuídas em suas mesorregiões⁷. O número entre parênteses se refere ao atribuído à localidade na rede de pontos do ALiB.

- Mesorregião Centro Norte Baiano – Itaberaba (090), Jacobina (086), Irecê (085);
- Mesorregião Centro Sul Baiano – Caetité (096), Itapetinga (100), Jequié (095), Seabra (089), Vitória da Conquista (098);
- Mesorregião Extremo Oeste Baiano – Barreiras (087), Santana (092);
- Mesorregião Metropolitana de Salvador – Capital – Salvador (093), Santo Amaro (091);
- Mesorregião Nordeste Baiano – Alagoinhas (088), Euclides da Cunha (083), Jeremoabo (082);
- Mesorregião Sul Baiano – Caravelas (102), Ilhéus (099), Santa Cruz de Cabrália (101), Valença (094);
- Mesorregião Vale São-Franciscano da Bahia – Barra (084), Carinhanha (097), Juazeiro (081).

⁷ Mesorregião. Unidade territorial homogênea, em nível maior que a microrregião, porém menor que o estado ou território, e resultado do grupamento de microrregiões. (Ferreira, 1986).

6.3. Localidades do Paraná

O conjunto de localidades do Paraná, que perfaz o total de 17 pontos, está distribuído em suas mesorregiões. O número entre parênteses se refere ao atribuído à localidade na rede de pontos do ALiB.

- Mesorregião Centro Ocidental Paranaense - Campo Mourão (212), Terra Boa (209);
- Mesorregião Centro Oriental Paranaense – Pirai do Sul (214);
- Mesorregião Centro-sul – Guarapuava (219); Mesorregião Metropolitana de Curitiba – Capital – Curitiba (220), Adrianópolis (216), Lapa (222), Morretes (221);
- Mesorregião Noroeste Paranaense – Nova Londrina (207), Umuarama (210);
- Mesorregião Norte Central Paranaense – Cândido de Abreu (213), Londrina (208);
- Mesorregião Norte Pioneiro Paranaense – Tomazina (211);
- Mesorregião Oeste Paranaense – São Miguel do Iguçu (217), Toledo (215);
- Mesorregião Sudeste Paranaense – Imbituva (218);
- Mesorregião Sudoeste Paranaense – Barracão (223).

6.4. Informantes

Em cada ponto de inquérito no interior dos dois estados foram entrevistados quatro informantes, dois homens e duas mulheres em duas faixas-etárias (18-30 anos e 50-65 anos). Nas capitais dos estados foram entrevistados oito informantes, quatro dos quais têm nível universitário. Informantes de 1 a 4 possuem nível fundamental e de 5 a 8 nível universitário. Os números ímpares se referem aos homens e os números pares se referem às mulheres; os números 1-2 e 5-6 são atribuídos aos informantes agrupados na primeira faixa etária e 3-4 e 7-8 à segunda faixa etária.

6.5. Questionário

No quadro 1, disposto em quatro colunas, apresentamos as duas questões do QSL utilizadas neste estudo.

Quadro 1- Extrato do QSL utilizado

QSL N°	Item Semântico-Lexical	Formulação da Pergunta	Áreas Semânticas
21	NEVOEIRO/CERRAÇÃO/ NEBLINA...	Muitas vezes, principalmente de manhã cedo, quase não se pode enxergar por causa de uma coisa parecida com fumaça, que cobre tudo. Como chamam isso?	FENÔMENOS ATMOSFÉRICOS
22	AMANHECER...	A parte do dia quando começa a clarear?	

7. Análise de dados

Neste capítulo, apresentamos a descrição e análise dos dados, estruturadas em itens que priorizam a perspectiva diatópica, mas contemplam, também, aspectos de natureza sociolinguística.

7.1. Salvador e Curitiba – a realidade das capitais

A questão 21 apura respostas para a pergunta: *Muitas vezes, principalmente de manhã cedo, quase não se pode enxergar por causa de uma coisa parecida com fumaça, que cobre tudo. Como chamam isso?* Houve 19 ocorrências das quais *neblina* teve maior produtividade, fornecida por seis informantes em Salvador e seis em Curitiba e corresponde a 63,3% das ocorrências. *Névoa* também não apresenta variação diatópica, pois há dois registros no ponto 093 e dois no ponto 220, logo significa 21% do total. As lexias que apresentam caráter diatópico são *nebulosidade* e *cerração*. *Nebulosidade* foi encontrada em Salvador e foi fornecida apenas por um informante e representa, portanto, 5,2% dentro das 19 ocorrências. *Cerração* foi fornecida por dois informantes em Curitiba e indica 10,5% do total.

A variante de maior produtividade para a questão 22 é *amanhecer* com nove ocorrências: quatro no ponto 093 e cinco no ponto 220. *Amanhecer* representa 69,5% das ocorrências. É necessário salientar que os informantes ao responder usaram flexões do verbo, como *está amanhecendo*, contudo, como indicado, consideramos o verbo no infinitivo. No nível da variação diatópica, temos *raiar do dia* com uma ocorrência em Salvador e representa 7,6% do total. Dois informantes em Salvador dizem que utilizam *alvorecer*. Para o informante 093/6 *alvorecer* responde a pergunta, entretanto o informante 093/7 usa *está alvorecendo* como segunda resposta. Sabemos que *alvorecer*

pode ser visto também como um substantivo, mas aqui o associamos ao verbo e consideramos o seu uso no particípio presente como resposta para a questão 22. *Alvorecer* representa 15,3% das ocorrências. Ainda no nível diatópico, registramos *alvorada* apenas em Curitiba com uma só ocorrência, o que representa 7,6% do total. Três informantes não souberam responder.

7.2. A realidade do interior dos dois estados

A questão 21 também nos traz 75 ocorrências e 9 abstenções. *Neblina* aparece como a variante mais produtiva nesta questão com 36 registros e indica 48% do total enquanto *neve* indica 37,4% com 28 ocorrências. As demais variantes para a questão 21 são *nevoeiro* com 4 registros: 5,3%; *cerração* com cinco registros: 6,7%; *nevoadão* e *névoa* cada uma com apenas um registro: 1,3% cada.

Houve 68 ocorrências e 16 abstenções para a questão 22. O número de variantes para esta questão é bastante significativo, com 12 lexias ou expressões distintas. A variante com maior número de registros é *amanhecer* que aparece 40 vezes no interior da Bahia e corresponde a 58,8% do total das ocorrências. A expressão *de manhã* ou simplesmente *manhã* aparece com sete registros representando 10,3% enquanto *clarear* apareceu com quatro registros significando 5,9% do total. *Madrugada* obteve quatro registros: 5,9% enquanto *barra do dia*, *manhãzinha* – ou *de manhãzinha*, *nascer* – o dia – receberam 2 registros cada: 2,9% cada. *Raiar do dia* aparece com uma ocorrência a mais do que as três últimas variantes, o que corresponde a 4,4% do total. As quatro expressões restantes, que são *início do dia*, *rompante do dia*, *alvorecer* e *romper do dia*, receberam um registro cada e cada uma delas representa 1,5% do total das ocorrências para a questão 22.

No Paraná, para a questão 21, *neblina* aparece com 34 registros representando 57,7% das 59 ocorrências. *Cerração* vem em segundo lugar com 24 registros correspondendo a 40,7%. *Nevoeiro* aparece com baixa produtividade com apenas um registro indicando 1,6% do total das ocorrências. Houve cinco abstenções na questão 21. A questão 22 apresenta 55 ocorrências e nove abstenções. *Amanhecer* apresenta 35 registros e representa a variante mais produtiva para esta questão. *Amanhecer* indica 63,7% das ocorrências. Em seguida temos *manhã* – ou *de manhã*. *Manhã* obteve 15 registros e representa 27,3%. *Manhãzinha* e *raiar do dia* competem em quantidade de

ocorrências uma vez que cada uma destas variantes obteve dois registros indicando 3,6% do total. *Aurora* obteve apenas um registro e indica 1,8%.

8. Uma visão plural entre a Bahia e o Paraná

Neste item, procuramos analisar de forma comparativa as duas áreas estudadas. O Quadro 2 mostra as lexias e expressões fraseológicas coincidentes nas cidades do interior da Bahia e Paraná, bem como lexias fornecidas apenas por informantes do interior da Bahia ou do interior do Paraná.

Quadro 2 – Coincidências e diferenças entre Bahia e Paraná

QUESTÕES	BAHIA/PARANÁ	BAHIA	PARANÁ
21	Neblina, nevoeiro, cerração	Neve, nevoado, névoa	_____
22	Amanhecer, raiar do dia, manhã, manhãzinha	Barra do dia, nascer o dia, início do dia, rompante do dia, alvorecer, romper do dia	Aurora

No quadro 2, temos o confronto da realidade lexical entre Bahia e Paraná. As lexias coincidentes na Bahia e no Paraná são aquelas geralmente de maior produtividade dentro das mesorregiões baianas e paranaenses. As formas que representam variação diatópica normalmente receberam um número menor de registros comparados com as variantes que não apresentam variação e estão distribuídas por todas as regiões de ambos os estados.

Como podemos mostrar nos gráficos 1 e 2:

Gráfico 1 – Neblina

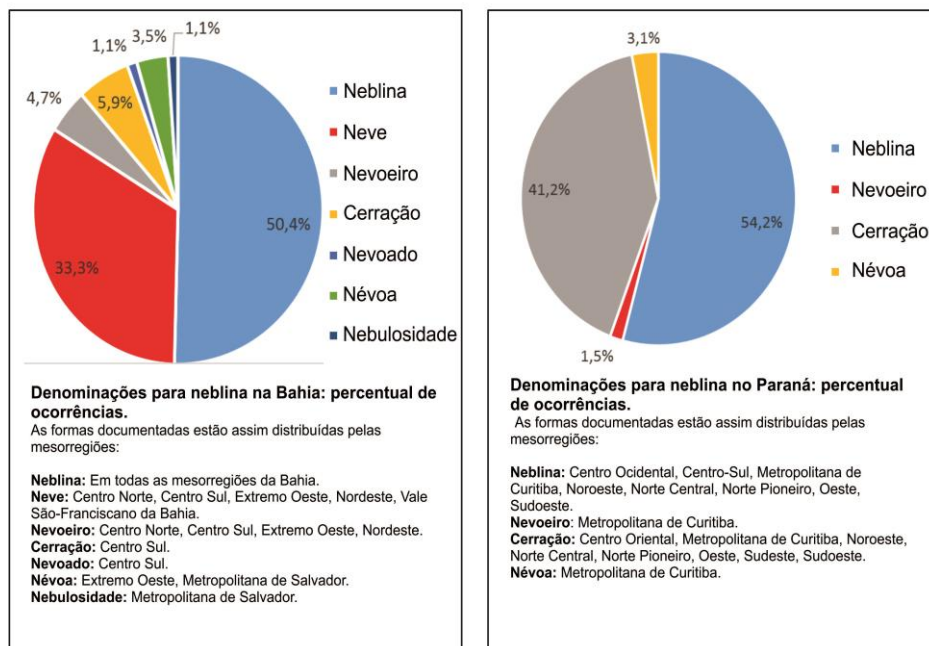
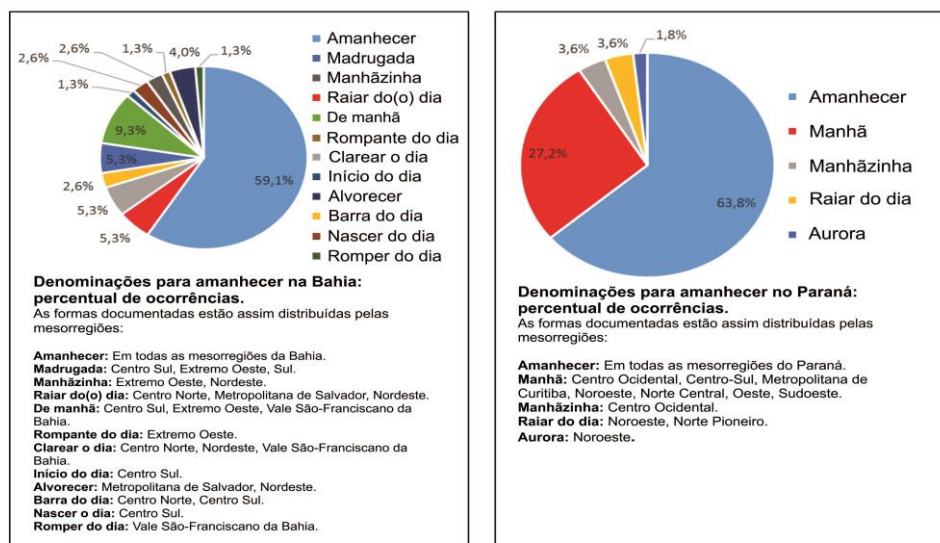


Gráfico 2 – Amanhecer



Gráficos 1 e 2: Fonte nossa. Dados do ALiB.

9. Variáveis sociais – um olhar sociolinguístico sobre o tema

Sereno foi fornecida por 093/1-2-3-4 e 220/1-3-4-6. Poderíamos dizer que *sereno* é uma variante diastrática fornecida por informantes com baixa escolaridade se não fosse pelo informante 220/6. Por outro lado, podemos observar, por conta destes dados, que a variante *sereno* é mais indicativa de baixa escolaridade, especialmente

porque o informante escolarizado que a forneceu não estava inicialmente seguro sobre sua resposta quando indagado. A seguir temos um diálogo para ratificar esta assertiva:

- *INQ.- De manhã cedo, a grama geralmente tá molhada. Como chamam aquilo que molha a grama?*
INF.- É... neblina?
INQ.- É? fica umas gotinhas assim na grama.
INF.- É ... ai é... como que é o negócio da noite que me falaram (risos). INQ.- Da noite.
INF.- Sereno.
INQ.- Ah...
INF.- É o sereno né que cai ali e fica molhado. (220/6).

Cerração, que aparece como uma variante diatópica, mostra-se também como uma variante diastrática e diageracional em Curitiba. Diastrática porque foi fornecida por um homem e uma mulher de nível fundamental. Diageracional porque este homem e esta mulher pertencem ao segundo grupo etário: 50 – 65 anos. *Cerração* parece ser uma variante mais conservadora que *neblina*. O informante 220/3 ilustra esta condição:

- *INQ.- Muitas vezes, principalmente de manhã cedo a gente quase não pode enxergar porque tem uma coisa que parece uma fumaça...*
INF.- Cerração. Cerração ou neblina, né.
INQ.- É o mesmo?
INF.- É o mesmo. A gente já... Antigamente também dizia cerração. INQ.- E agora o senhor acha que...?
INF.- Atualmente é chamado de neblina, né. “Tá uma neblina forte.” Antigamente não. “Oh, caiu uma cerração forte aí”.
INQ.- Mudou um pouco né o jeito de falar. (220/3).

Para as anotações referentes à pergunta 22, constatamos que:

- *Amanhecer, nascer o sol, pôr do sol, anoitecer, Estrela D'alva, estrela cadente, cair* não apresentam variação social.

É importante salientar que as conclusões de cunho sociolinguístico não podem ser categóricas por conta do número pequeno de informantes arrolados para esta pesquisa.

Considerações finais

Neste trabalho identificamos, analisamos e descrevemos a variação linguística encontrada nos estados brasileiros, Bahia e Paraná, localizados, respectivamente, nas regiões Nordeste e Sul do Brasil, com vistas a mostrar a homogeneidade ou

heterogeneidade desses dois falares brasileiros no que diz respeito à área semântica *fenômenos atmosféricos*.

Para as questões 21 e 22, as variantes mais produtivas *neblina* e *amanhecer* aparecem distribuídas por todas as mesorregiões da Bahia e do Paraná. As variantes que ocorreram apenas na Bahia ou apenas no Paraná não foram fornecidas por um número expressivo de informantes.

Considerando as questões iniciais, o léxico documentado na Bahia e Paraná oferece características particulares que podem delimitar áreas e refletir a natureza da constituição histórica de cada um desses estados. Contudo, verificamos que no plano da análise diatópica, podemos destacar nesse recorte analisado, como resultado mais relevante o fato de a distribuição das variantes terem um caráter mais homogêneo do que heterogêneo considerando os dados de Salvador e Curitiba bem como de todas as cidades do interior dos dois estados.

Referências

AGUILERA, Vanderci de Andrade. *Atlas Linguístico do Paraná. Vol. I*. Tese apresentada à Faculdade de Ciências e Letras da Universidade Estadual Paulista – UNESP, Assis, 1990.

AGUILERA, Vanderci de Andrade. *Tupinismos lexicais no português brasileiro: trilhas e traços no Paraná*. Sonderdruck aus Ex oriente lux – Festschrift für Eberhard Gartner zu seinem 60. Geburtstag. Herausgegeben von Sybille GroBe und Axel Schonberger in Verbindung mit Cornelia Doll und Christine Hundt. Valentia: Frankfurt am Main, 2002.

AMARAL, Amadeu. *O Dialecto Caipira. Gramática- Vocabulário*. São Paulo: O Livro, 1920.

BAGNO, Marcos. *Preconceito Linguístico – O que é, como se faz*. São Paulo: Loyola, 2000.

CARDOSO, Suzana Alice Marcelino. O Atlas Linguístico do Brasil: um projeto nacional. In: AGUILERA, Vanderci de A. (Org.). *A Geolinguística no Brasil – caminhos e perspectivas*. Londrina: UEL, 1998.

CARDOSO, Suzana Alice Marcelino. Caminhos da Dialectologia Brasileira. In: ALVES, Eliane Ferraz; ESPÍNDOLA, Luciene C; DA HORA, Dermeval. (Orgs.) *ABRALIN: 40 anos em cena*. João Pessoa: CCHLA, 2009.

CARDOSO, Suzana Alice Marcelino. *Geolinguística – Tradição e Modernidade*. São Paulo: Parábola, 2010.

- CHAMBERS, J.K.; TRUDGILL, Peter. *La Dialectologia*. Traducción Carmen Morán González. Madrid: Visor Libros, 1994.
- COSERIU, Eugenio. *Sincronia, Diacronia e História – O problema da mudança linguística*. Tradução de Carlos Alberto da Fonseca e Mário Ferreira. Rio de Janeiro: Presença, 1979.
- COSTA, F. A. Pereira da. *O vocabulário pernambucano*. Recife: Secretaria de Educação e Cultura; MEC, 1976.
- FERREIRA, Aurélio Buarque de Holanda. *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro: Nova Fronteira, 1986.
- FERREIRA, Carlota; CARDOSO, Suzana Alice. *A Dialectologia no Brasil*. São Paulo: Contexto, 1994.
- FINCH, Geoffrey. *Linguistics Terms and Concepts*. New York: St. Martin's Press, 2000.
- LABOV, William. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008.
- MORENO FERNÁNDEZ, Francisco. *Principios de Sociolingüística y sociología del lenguaje*. Barcelona: Editorial Ariel, 1998.
- MOTA, Jacyra Andrade. Áreas Dialectais Brasileiras. In: *Quinhentos Anos de História Linguística do Brasil*. Orgs. CARDOSO, Suzana Alice M.; MOTA, Jacyra Andrade; MATTOS E SILVA, Rosa Virgínia. Salvador: Apoio, 2006.
- NASCENTES, Antenor. *O Linguajar Carioca*. 2ª Ed. Rio de Janeiro: Organizações Simões, [1922], 1953.
- PRETI, Dino. *Sociolingüística – os níveis da fala*. São Paulo: Edusp, 2003.
- ROSSI, Nelson; FERREIRA, Carlota; ISENSEE, Dinah. *Atlas Prévio dos Falares Baianos*. Rio de Janeiro: Ministério de Educação e Cultura; Instituto Nacional do Livro, 1963.
- QSL. *Questionário Semântico-Lexical*. Comitê Nacional do Projeto ALiB – Atlas Linguístico do Brasil. Questionários 2001. Londrina: UEL, 2001.
- REIS, João José. *Rebelião Escrava no Brasil – A História do Levante dos Malês em 1835*. Edição revista e ampliada. São Paulo: Companhia das Letras, 2009.